Distance between metric measure spaces and distance matrix distributions

著者（英）	Ryunosuke OZAWA
journal or publication title	Tsukuba Journal of Mathematics
volume	38
number	2
page range	159-170
year	2014-10
URL	http://hdl.handle.net/2241/00159025

TSUKUBA J. MATH.
Vol. 38 No. 2 (2014), 159–170

DISTANCE BETWEEN METRIC MEASURE SPACES AND DISTANCE MATRIX DISTRIBUTIONS

By

Ryunosuke Ozawa

Abstract. We study the Prohorov distance between the distance matrix distributions of two metric measure spaces. We prove that it is not smaller than 1-box distance between two metric measure spaces and also prove that it is not larger than 0-box distance between two metric measure spaces.

1. Introduction

In this paper, we consider the relation between the box distance function and the distance matrix distribution. Here, a metric measure space $X := (X, d_X, \mu_X)$ is a complete separable metric space (X, d_X) equipped with a Borel probability measure μ_X. The box distance function $\underline{\square}_\lambda$, which was introduced by Gromov [5], is a natural distance function between two metric measure spaces for any $\lambda \geq 0$. Define the distance map K_r^X from X^r to $M_r(\mathbf{R})$, the set of all real square matrices of order r with l_∞-norm $\|\cdot\|_\infty$, by $K_r^X(x_1, \ldots, x_r) := (d_X(x_k, x_l))_{k,l=1,\ldots,r}$. The r-dimensional distance matrix distribution $\underline{\mu}_r^X$ is the push-forward measure of the r-times product measure of μ_X by K_r^X. $\underline{\mu}_r^X$ is a Borel probability measure on $M_r(\mathbf{R})$. We denote by $\mathscr{X}$ the set of isomorphism classes of metric measure spaces. Gromov [5] developed a theory of convergence of metric measure spaces in $\mathscr{X}$. He proved that the infinite-dimensional distance matrix distribution $\underline{\mu}_\infty^X$ is a complete invariant of metric measure spaces (see Theorem 2.3). On the other hand, Greven-Pfaffelhuber-Winter [4] developed a theory of convergence of metric measure spaces and the probabilistic theory on $\mathscr{X}$, independently of Gromov's

---

2010 *Mathematics Subject Classification*: Primary 53C23, 28D05.
Key words and phrases: mm-space, box distance function, distance matrix distribution.
This work was partially supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.
Received July 16, 2013.
Revised September 3, 2014.

work. They introduced the Gromov-Prohorov metric d_{GPr} and proved that the topology of d_{GPr} is compatible with the weak convergence of infinite-dimensional distance matrix distributions (see Theorem 2.12). After that, Löhr [7] showed that $\underline{\square}_\lambda$ ($\lambda > 0$) and d_{GPr} are bi-Lipschitz equivalent to each other (see Theorem 2.16). Thus, the topology of the box distance function $\underline{\square}_\lambda$ ($\lambda > 0$) is compatible with the weak convergence of infinite-dimensional distance matrix distributions.

Our purpose is to estimate the box distance between two metric measure spaces X and Y by its the distance matrix distributions $\underline{\mu}_\infty^X$ and $\underline{\mu}_\infty^Y$. In this paper, we consider the distance between two distance matrix distributions $\underline{\mu}_\infty^X$ and $\underline{\mu}_\infty^Y$ of two metric measure spaces X and Y. Moreover, we interpret it as a new metric on $\mathcal{X}$.

DEFINITION 1.1. Let $X, Y \in \mathcal{X}$. We define the distance $d_{l_\infty\text{-}\mathrm{Pr}}(X, Y)$ to be the Prohorov distance between $\underline{\mu}_\infty^X$ and $\underline{\mu}_\infty^Y$ with respect to the l_∞-norm on $M_\infty(\mathbf{R})$.

REMARK 1.2. (1) Symmetry and triangle inequality of Prohorov metric (see Proposition 2.5) and Theorem 2.3 imply that $d_{l_\infty\text{-}\mathrm{Pr}}$ is a metric on $\mathcal{X}$.
(2) The l_∞-norm $\|\cdot\|_\infty$ may take the value ∞ on $M_\infty(\mathbf{R})$ but the set of Borel probability measures on $M_\infty(\mathbf{R})$ equipped with the Prohorov metric with respect to the l_∞-norm is a metric space.
(3) $d_{l_\infty\text{-}\mathrm{Pr}}(X, Y) \leq 1$ for any two mm-spaces X and Y.

We study the relation between the box distance $\underline{\square}_\lambda(X, Y)$ and $d_{l_\infty\text{-}\mathrm{Pr}}(X, Y)$. Our main result is stated as follows.

THEOREM 1.3. *Let $X, Y \in \mathcal{X}$. We have*

$$\underline{\square}_1(X, Y) \leq d_{l_\infty\text{-}\mathrm{Pr}}(X, Y) \leq \underline{\square}_0(X, Y).$$

We next compare the topologies induced from $\underline{\square}_1$ and $d_{l_\infty\text{-}\mathrm{Pr}}$.

PROPOSITION 1.4. *Let $X := (\{p_1\}, \delta_{p_1})$ and $X_n := (\{p_1, p_2\}, d_{\{p_1,p_2\}}, (1 - n^{-1})\delta_{p_1} + n^{-1}\delta_{p_2})$, where $d_{\{p_1,p_2\}}(p_1, p_2) := 1$. We denote by δ_{p_i} the Dirac measure at p_i $(i = 1, 2)$. Then $\underline{\square}_1(X_n, X) = n^{-1}$ but $d_{l_\infty\text{-}\mathrm{Pr}}(X_n, X) = \underline{\square}_0(X_n, X) = 1$ for any $n \geq 2$.*

This proposition means that $\underline{\mu}_\infty^{X_n}$ converges to $\underline{\mu}_\infty^X$ weakly as $n \to \infty$ but the Prohorov distance between $\underline{\mu}_\infty^{X_n}$ and $\underline{\mu}_\infty^X$ is one for any $n \geq 2$ and implies the next corollary.

Corollary 1.5. *$(\mathscr{X}, d_{l_\infty\text{-Pr}})$ and $(\mathscr{X}, \underline{\square}_1)$ are not homeomorphic to each other by the identity map.*

Corollary 1.5 seems not to fit to the fact that the topology of the Prohorov metric is compatible with the weak convergence of probability measures on the complete separable metric space (see Theorem 2.7). Since $(M_\infty(\mathbf{R}), \|\cdot\|_\infty)$ is not separable, Corollary 1.5 does not contradictory. We do not know if the topology generated by $\underline{\square}_0$ coincides with the topology induced from $d_{l_\infty\text{-Pr}}$ or not.

2. Preliminaries

2.1. Metric Measure Space. Let $(X, \mathcal{O}_X)$ be a topological space. We denote by $\mathscr{B}(X)$ the Borel σ-algebra and $\mathscr{M}(X)$ the set of all Borel probability measures on X. Let μ be a Borel probability measure on X. Recall that the *support of* μ, $\mathrm{supp}(\mu)$, is the smallest closed set $\mathrm{supp}(\mu) \subset X$ such that $\mu(X \backslash \mathrm{supp}(\mu)) = 0$. The *push forward of μ by a measurable map φ from X into another topological space* $(Y, \mathcal{O}_Y)$ is the Borel probability measure $\varphi_*\mu \in \mathscr{M}(Y)$ defined by

$$\varphi_*\mu(A) := \mu(\varphi^{-1}(A)),$$

for all $A \in \mathscr{B}(Y)$.

Definition 2.1 (Metric measure space). A triple $X := (X, d_X, \mu_X)$ is called a *metric measure space* (*or mm-space*) if (X, d_X) is a complete separable metric space and if μ_X is a Borel probability measure on X. Two mm-spaces (X, d_X, μ_X) and (Y, d_Y, μ_Y) are *isomorphic* if there exists an isometry φ between the supports of μ_X on (X, d_X) and of μ_Y on (Y, d_Y) such that $\mu_Y = \varphi_*\mu_X$. We write $\mathscr{X}$ by the set of isomorphism classes of mm-spaces.

Let $M_\infty(\mathbf{R})$ be the set of all real square matrices of infinite order equipped with the coarsest topology such that the natural projection $\mathrm{pr}_{\infty,r} : M_\infty(\mathbf{R}) \to M_r(\mathbf{R})$ is continuous for any $r \in \mathbf{R}$.

Definition 2.2 (Distance matrix distribution). Let $X = (X, d_X, \mu_X) \in \mathscr{X}$ and $r \in \mathbf{N} \cup \{\infty\}$. Define a map $K_r^X : X^r \to M_r(\mathbf{R})$ by

$$K_r^X(x_1, \ldots, x_r) := (d_X(x_k, x_l))_{k,l=1,\ldots,r},$$

and the *r-dimensional distance matrix distribution* $\underline{\mu}_r^X$ *of* X by

$$\underline{\mu}_r^X := (K_r^X)_* \mu_X^{\otimes r},$$

where $\mu_X^{\otimes r}$ is the r times product measure of μ_X.

Theorem 2.3 (mm-Reconstruction theorem, [5, Section $3\frac{1}{2}.5$, Section $3\frac{1}{2}.7$], [8, Section 2, Theorem], [6, Theorem 2.1]). *Let $X, Y \in \mathcal{X}$. The following* (1), (2), *and* (3) *are equivalent to each other.*

(1) *X and Y are isomorphic to each other.*

(2) $\underline{\mu}_r^X = \underline{\mu}_r^Y$ *for all* $r \in \mathbf{N}$.

(3) $\underline{\mu}_\infty^X = \underline{\mu}_\infty^Y$.

This theorem means that the infinite-dimensional distance matrix distribution is a complete invariant of mm-spaces.

2.2. Gromov-Prohorov Metric. Let (X, d_X) be a metric space. For a real number $r > 0$ and a subset $A \subset X$, we set $B_r(A) := \{x \in X \mid d_X(x, A) < r\}$, where $d_X(x, A) := \inf_{x' \in A} d_X(x, x')$.

Definition 2.4 (Prohorov metric). Define the *Prohorov metric* $d_{\mathrm{Pr}}^{(X, d_X)}$ *on* $\mathcal{M}(X)$ by

$$d_{\mathrm{Pr}}^{(X, d_X)}(\mu, \nu) := \inf\{\varepsilon > 0 \mid \mu(A) \le \nu(B_\varepsilon(A)) + \varepsilon, \ \textit{for all} \ A \in \mathcal{B}(X)\}$$

for $\mu, \nu \in \mathcal{M}(X)$.

Note that $d_{\mathrm{Pr}}^{(X, d_X)}(\mu, \nu) \le 1$ for any two Borel probability measures μ and ν on X.

Proposition 2.5 ([2, Lemma 3.1.1]). $(\mathcal{M}(X), d_{\mathrm{Pr}}^{(X, d_X)})$ *is a metric space.*

Definition 2.6 (Weak convergence). We say that a sequence $\{\mu_n\}_{n=1}^\infty$ *of Borel probability measures on X converges weakly to a Borel probability measure μ on X* and write $\mu_n \to \mu$ *weakly as* $n \to \infty$ if

$$\lim_{n \to \infty} \int_X f(x) \, d\mu_n = \int_X f(x) \, d\mu$$

for any bounded continuous function $f : X \to \mathbf{R}$.

Theorem 2.7 ([2, Theorem 3.3.1]). *Let (X, d_X) be a separable metric space, $\{\mu_n\}_{n=1}^{\infty}$ a sequence of Borel probability measures on X, and μ a Borel probability measure on X. Then*

$$\lim_{n\to\infty} d_{\mathrm{Pr}}^{(X,d_X)}(\mu_n, \mu) = 0$$

if and only if

$$\mu_n \to \mu \text{ weakly as } n \to \infty.$$

For $\mu, \nu \in \mathcal{M}(X)$, we say $\mu \leq \nu$ if $\mu(A) \leq \nu(A)$ for any $A \in \mathcal{B}(X)$. A finite Borel measure π on $X \times X$ is called a *partial transport plan from $\mu \in \mathcal{M}(X)$ to $\nu \in \mathcal{M}(X)$* if $(\mathrm{pr}_1)_*\pi \leq \mu$ and $(\mathrm{pr}_2)_*\pi \leq \nu$, where $\mathrm{pr}_i : X \times X \to X$, $i = 1, 2$, are the projections defined by $\mathrm{pr}_1(x, x') = x$, $\mathrm{pr}_2(x, x') = x'$. For a partial transport plan π from μ to ν, we define the *deficiency* $\operatorname{def} \pi$ *of* π by $\operatorname{def} \pi := 1 - \pi(X \times X)$. For $\varepsilon \geq 0$, the partial transport plan π is called an *ε-transport plan from μ to ν* if

$$\mathrm{supp}(\pi) \subset X(\varepsilon) := \{(x, x') \in X \times X \mid d_X(x, x') \leq \varepsilon\}.$$

Definition 2.8 (Transportation distance). Let $\mu, \nu \in \mathcal{M}(X)$. Define the *Transportation distance* $\mathrm{Tra}^{(X,d_X)}$ *between μ and ν* by

$$\mathrm{Tra}^{(X,d_X)}(\mu, \nu) := \inf\{\varepsilon > 0 \mid \textit{there exists an } \varepsilon\textit{-transport plan } \pi \textit{ from } \mu \textit{ to } \nu \textit{ satisfying } \operatorname{def} \pi \leq \varepsilon\}.$$

Theorem 2.9 (Strassen's theorem, [9, Corollary 1.28]). *Let (X, d_X) be a complete separable metric space. For any $\mu, \nu \in \mathcal{M}(X)$, we have*

$$d_{\mathrm{Pr}}^{(X,d_X)}(\mu, \nu) = \mathrm{Tra}^{(X,d_X)}(\mu, \nu).$$

Next, we define the Gromov-Prohorov metric on $\mathcal{X}$.

Definition 2.10 (Gromov-Prohorov metric, [4]). Let $X = (X, d_X, \mu_X)$, $Y = (Y, d_Y, \mu_Y)$ be two mm-spaces. Define the *Gromov-Prohorov metric d_{GPr} on $\mathcal{X}$* by

$$d_{\mathrm{GPr}}(X, Y) := \inf_{(\varphi_X, \varphi_Y, Z)} d_{\mathrm{Pr}}^{(Z,d_Z)}((\varphi_X)_*\mu_X, (\varphi_Y)_*\mu_Y),$$

where the infimum is taken over all isometric embeddings φ_X and φ_Y from $\mathrm{supp}(\mu_X)$ and $\mathrm{supp}(\mu_Y)$, respectively, into some common metric space (Z, d_Z).

Note that $d_{\mathrm{GPr}}(X, Y) \le 1$ for any two mm-spaces X and Y.

THEOREM 2.11 ([4, Theorem 1]). *$(\mathcal{X}, d_{\mathrm{GPr}})$ is a metric space.*

THEOREM 2.12 ([4, Theorem 5]). *Let $X, X_1, X_2, \ldots \in \mathcal{X}$. Then*

$$\lim_{n\to\infty} d_{\mathrm{GPr}}(X_n, X) = 0$$

if and only if

$$\underline{\mu}_{\infty}^{X_n} \to \underline{\mu}_{\infty}^{X} \text{ weakly as } n \to \infty.$$

2.3. Gromov's Box Distance. We denote by $\mathcal{L}$ the Lebesgue measure on $[0,1]$. For any mm-space X there exists a Borel measurable map $p_X : [0,1] \to X$ with $(p_X)_*\mathcal{L} = \mu_X$ (see [1, Theorem 9.4.7]). We call such a map p_X a *parameter of* X. Note that a parameter of X is not unique in general. For a parameter p_X of X, we define a function $(p_X)^* d_X : [0,1] \times [0,1] \to \mathbf{R}$ by $(p_X)^* d_X(s, s') := d_X(p_X(s), p_X(s'))$.

DEFINITION 2.13 (Box distance, [5]). Let $\lambda \ge 0$, and $X = (X, d_X, \mu_X)$, $Y = (Y, d_Y, \mu_Y) \in \mathcal{X}$. Define the *box distance between X and Y* by

$$\begin{aligned}\square_\lambda(X, Y) := \inf_{(p_X, p_Y)} \{\varepsilon > 0 \mid & \textit{ there exists } T_\varepsilon \in \mathcal{B}([0,1]) \textit{ such that } \mathcal{L}(T_\varepsilon) \ge 1 - \lambda\varepsilon \\ & \textit{ and } |(p_X)^* d_X(s, s') - (p_Y)^* d_Y(s, s')| \le \varepsilon \textit{ for all } s, s' \in T_\varepsilon\},\end{aligned}$$

where the infimum is taken over all parameters $p_X : [0,1] \to X$ and $p_Y : [0,1] \to Y$.

Note that $\square_\lambda(X, Y) \le 1/\lambda$ for any two mm-spaces X, Y, and $\lambda > 0$.

THEOREM 2.14 ([5, Section $3\frac{1}{2}$.6], [3, Theorem 3.1.8]). *$(\mathcal{X}, \square_\lambda)$ is a metric space for any $\lambda \ge 0$.*

It is easy to see that $\square_\lambda$ for all $\lambda > 0$ are bi-Lipschitz equivalent to each other and $\square_1 \le \square_0$.

PROPOSITION 2.15 ([5, Section $3\frac{1}{2}$.10]). *Let (X, d_X) be a complete separable metric space and $\mu, \nu \in \mathcal{M}(X)$. Then we have*

$$\square_1((X, d_X, \mu), (X, d_X, \nu)) \le d_{\mathrm{Pr}}^{(X, d_X)}(\mu, \nu).$$

Finally, we consider the relation between the Gromov-Prohorov metric and the box distance function.

THEOREM 2.16 ([7, Corollary 6]). *For any* $X, Y \in \mathscr{X}$, *we have*

$$d_{\mathrm{GPr}}(X, Y) \leq \underline{\square}_1(X, Y) \leq 2d_{\mathrm{GPr}}(X, Y). \tag{2.1}$$

Moreover, $(\mathscr{X}, d_{\mathrm{GPr}})$ *and* $(\mathscr{X}, \underline{\square}_1)$ *are homeomorphic to each other.*

REMARK 2.17. (1) Löhr [7] proved that $d_{\mathrm{GPr}} = \frac{1}{2}\underline{\square}_{1/2}$ on $\mathscr{X}$. This implies the inequality (2.1).
(2) Theorem 2.12 and Theorem 2.16 imply that X_n $\underline{\square}_1$-converges X as $n \to \infty$ if and only if $\underline{\mu}_\infty^{X_n}$ converses weakly to $\underline{\mu}_\infty^X$ as $n \to \infty$. The proof of this statement is omitted in the original article (see [5, Section $3\frac{1}{2}$.14]).

3. Box Distance and Distance Matrix Distribution

In this section, we give the proof of Theorem 1.3.

Let $r, r' \in \mathbf{N} \cup \{\infty\}$ with $r \geq r'$. Define the projection $\mathrm{pr}_{r,r'} : M_r(\mathbf{R}) \to M_{r'}(\mathbf{R})$ by

$$\mathrm{pr}_{r,r'}((a_{kl})_{k,l=1,\dots,r}) := (a_{kl})_{k,l=1,\dots,r'}.$$

LEMMA 3.1 ([6, Lemma 2.2]). *Let* $X \in \mathscr{X}$ *and* $r \in \mathbf{N}$. *We have* $(\mathrm{pr}_{r+1,r})_* \underline{\mu}_{r+1}^X = \underline{\mu}_r^X$ *and* $(\mathrm{pr}_{\infty,r})_* \underline{\mu}_\infty^X = \underline{\mu}_r^X$.

LEMMA 3.2. *Let* $X, Y \in \mathscr{X}$. *Then* $d_{\mathrm{Pr}}^{(M_r(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y)$ *is monotone non-decreasing in* $r \in \mathbf{N}$. *In particular,*

$$\sup_{r \in \mathbf{N}} d_{\mathrm{Pr}}^{(M_r(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) = \lim_{r \to \infty} d_{\mathrm{Pr}}^{(M_r(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y).$$

PROOF. Let $\varepsilon > 0$ satisfy $d_{\mathrm{Pr}}^{(M_{r+1}(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_{r+1}^X, \underline{\mu}_{r+1}^Y) < \varepsilon$. By the definition of Prohorov metric, we get $\underline{\mu}_{r+1}^X(A) \leq \underline{\mu}_{r+1}^Y(B_\varepsilon(A)) + \varepsilon$ for all $A \in \mathscr{B}(M_{r+1}(\mathbf{R}))$. Since $\mathrm{pr}_{r+1,r}^{-1}(A') \in \mathscr{B}(M_{r+1}(\mathbf{R}))$ for any $A' \in \mathscr{B}(M_r(\mathbf{R}))$, we have

$$\underline{\mu}_{r+1}^X(\mathrm{pr}_{r+1,r}^{-1}(A')) \leq \underline{\mu}_{r+1}^Y(B_\varepsilon(\mathrm{pr}_{r+1,r}^{-1}(A'))) + \varepsilon.$$

Obviously, $B_\varepsilon(\mathrm{pr}_{r+1,r}^{-1}(A')) = \mathrm{pr}_{r+1,r}^{-1}(B_\varepsilon(A'))$. Therefore,

$$\underline{\mu}_{r+1}^X(\mathrm{pr}_{r+1,r}^{-1}(A')) \leq \underline{\mu}_{r+1}^Y(\mathrm{pr}_{r+1,r}^{-1}(B_\varepsilon(A'))) + \varepsilon.$$

By Lemma 3.1, this means

$$\underline{\mu}_r^X(A') \leq \underline{\mu}_r^Y(B_\varepsilon(A')) + \varepsilon,$$

then we obtain $d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) \leq \varepsilon$. By the arbitrariness of ε, we have $d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) \leq d_{\mathrm{Pr}}^{(M_{r+1}(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_{r+1}^X, \underline{\mu}_{r+1}^Y)$. □

LEMMA 3.3. *Let $X, Y \in \mathscr{X}$. Then we have*

$$d_{l_\infty\text{-Pr}}(X, Y) = \lim_{r\to\infty} d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y).$$

PROOF. The inequality

$$\lim_{r\to\infty} d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) \leq d_{l_\infty\text{-Pr}}(X, Y) \tag{3.1}$$

is obtained in the same way as in the proof of Lemma 3.2.

Next, we prove the inequality

$$d_{l_\infty\text{-Pr}}(X, Y) \leq \sup_{r\in\mathbf{N}} d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) \tag{3.2}$$

by Lemma 3.2. Let $\varepsilon' > 0$ satisfy $d_{\mathrm{Pr}}^{(M_r(\mathbf{R}),\|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) < \varepsilon'$ for any $r \in \mathbf{N}$, and $A' \in \mathscr{B}(M_\infty(\mathbf{R}))$. By the definition of Prohorov metric, Lemma 3.1 and $\mathrm{pr}_{\infty,r}^{-1}(B_{\varepsilon'}(A)) = B_{\varepsilon'}(\mathrm{pr}_{\infty,r}^{-1}(A))$ for any $A \in \mathscr{B}(M_r(\mathbf{R}))$, we get

$$\underline{\mu}_\infty^X(\mathrm{pr}_{\infty,r}^{-1}(\mathrm{pr}_{\infty,r}(A'))) \leq \underline{\mu}_\infty^Y(B_{\varepsilon'}(\mathrm{pr}_{\infty,r}^{-1}(\mathrm{pr}_{\infty,r}(A')))) + \varepsilon'.$$

Using the continuity of measure for $\{\bigcap_{n=1}^r \mathrm{pr}_{\infty,n}^{-1}(\mathrm{pr}_{\infty,n}(A'))\}_{r=1}^\infty$, we have

$$\lim_{r\to\infty} \underline{\mu}_\infty^X(\mathrm{pr}_{\infty,r}^{-1}(\mathrm{pr}_{\infty,r}(A'))) = \underline{\mu}_\infty^X(A') \tag{3.3}$$

and

$$\lim_{r\to\infty} \underline{\mu}_\infty^Y(B_{\varepsilon'}(\mathrm{pr}_{\infty,r}^{-1}(\mathrm{pr}_{\infty,r}(A')))) = \underline{\mu}_\infty^Y(B_{\varepsilon'}(A')). \tag{3.4}$$

(3.3) and (3.4) lead to $\underline{\mu}_\infty^X(A') \leq \underline{\mu}_\infty^Y(B_{\varepsilon'}(A')) + \varepsilon'$. Then this means

$$d_{l_\infty\text{-Pr}}(X, Y) \leq \varepsilon'.$$

We obtain (3.2).

Combining two inequalities (3.1) and (3.2), we have the lemma. □

To prove Theorem 1.3, we need a uniformly distributed sequence.

Definition 3.4 (Uniformly distributed sequence). Let $X \in \mathcal{X}$ and $\{x_i\}_{i=1}^{\infty} \subset X$. $\{x_i\}_{i=1}^{\infty}$ is called a *uniformly distributed sequence of* X if

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i} \to \mu_X \text{ weakly as } n \to \infty,$$

where δ_{x_i} is the Dirac measure at x_i. We write E_X by the set of all uniformly distributed sequences of X.

Consider E_X as a subset of $X^{\mathbf{N}}$. The next lemma means that there are many uniformly distributed sequences of X.

Lemma 3.5 ([6, Lemma 2.4]). *Let $X \in \mathcal{X}$. We have $\mu_X^{\otimes \mathbf{N}}(E_X) = 1$ and in particular, $\underline{\mu}_{\infty}^{X}(K_{\infty}^{X}(E_X)) = 1$.*

Finally, we prove the main theorem.

Proof of Theorem 1.3. We first prove the inequality $\underline{\square}_1(X, Y) \le d_{l_\infty\text{-Pr}}(X, Y)$. This is trivial in the case of $d_{l_\infty\text{-Pr}}(X, Y) = 1$. Let $0 < \varepsilon < 1$ satisfy $d_{l_\infty\text{-Pr}}(X, Y) < \varepsilon$. We apply the definition of $d_{l_\infty\text{-Pr}}$ for $A = K_{\infty}^{X}(E_X)$ and use Lemma 3.5 to have

$$\underline{\mu}_{\infty}^{Y}(B_\varepsilon(K_{\infty}^{X}(E_X))) \ge 1 - \varepsilon > 0.$$

By $\underline{\mu}_{\infty}^{Y}(K_{\infty}^{Y}(E_Y)) = 1$, we have $B_\varepsilon(K_{\infty}^{X}(E_X)) \cap K_{\infty}^{X}(E_Y) \neq \emptyset$. Then there exist two sequences $\{x_i\}_{i=1}^{\infty} \in E_X$ and $\{y_i\}_{i=1}^{\infty} \in E_Y$ such that

$$|d_X(x_i, x_j) - d_Y(y_i, y_j)| < \varepsilon$$

for all $i, j \in \mathbf{N}$. Define mm-spaces X_n and Y_n by

$$X_n := \left(X, d_X, \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}\right), \quad Y_n := \left(Y, d_Y, \frac{1}{n}\sum_{i=1}^{n} \delta_{y_i}\right).$$

From the definition of uniformly distributed sequence and Proposition 2.15, for any $\delta > 0$, there exists a number $n_0 \in \mathbf{N}$ such that $\underline{\square}_1(X_n, X) < \delta$ and $\underline{\square}_1(Y_n, Y) < \delta$ for any $n \ge n_0$. Define two parameters p_{X_n} and p_{Y_n} of X_n and Y_n by

$$p_{X_n}(s) := \begin{cases} x_i & \text{if } s \in \left[\dfrac{i-1}{n}, \dfrac{i}{n}\right),\ i = 1, \ldots, n, \\ x_n & \text{if } s = 1, \end{cases}$$

$$p_{Y_n}(s) := \begin{cases} y_i & \text{if } s \in \left[\dfrac{i-1}{n}, \dfrac{i}{n}\right),\ i = 1, \ldots, n, \\ y_n & \text{if } s = 1. \end{cases}$$

We then have

$$|(p_{X_n})^* d_X(s,s') - (p_{Y_n})^* d_Y(s,s')| \le \varepsilon$$

for any $s, s' \in [0,1]$. This implies $\underline{\square}_1(X_n, Y_n) < \varepsilon$. Thus, we have

$$\begin{aligned}\underline{\square}_1(X, Y) &\le \underline{\square}_1(X, X_n) + \underline{\square}_1(X_n, Y_n) + \underline{\square}_1(Y_n, Y) \\ &< 2\delta + \varepsilon.\end{aligned}$$

By the arbitrariness of δ and ε, we have $\underline{\square}_1(X, Y) \le d_{l_\infty\text{-Pr}}(X, Y)$.

We next prove the inequality $d_{l_\infty\text{-Pr}}(X, Y) \le \underline{\square}_0(X, Y)$. Let $\varepsilon > 0$ satisfy $\underline{\square}_0(X, Y) < \varepsilon$. From the definition of box distance, there exist two parameters p_X, p_Y of X, Y and a Borel set T_ε on $[0,1]$ such that $\mathscr{L}(T_\varepsilon) = 1$ and

$$|(p_X)^* d_X(s,s') - (p_Y)^* d_Y(s,s')| < \varepsilon$$

for all $s, s' \in T_\varepsilon$. Define the map $\xi_r : (T_\varepsilon)^r \to M_r(\mathbf{R}) \times M_r(\mathbf{R})$ by

$$\xi_r(s_1, \dots, s_r) := (K_r^X(p_X(s_1), \dots, p_X(s_r)), K_r^Y(p_Y(s_1), \dots, p_Y(s_r))).$$

We set $\pi_r := (\xi_r)_* \mathscr{L}^{\otimes r}$. This belongs to $\mathscr{M}(M_r(\mathbf{R}) \times M_r(\mathbf{R}))$. We will prove that π_r is an ε-transportation from $\underline{\mu}_r^X$ to $\underline{\mu}_r^Y$ and $\operatorname{def} \pi_r \le \varepsilon$. Obviously,

$$\operatorname{supp}(\pi_r) \subset \xi_r((T_\varepsilon)^r) \subset M_r(\mathbf{R})(\varepsilon).$$

Define two projections $\mathrm{pr}_1 : M_r(\mathbf{R}) \times M_r(\mathbf{R}) \to M_r(\mathbf{R})$, $\mathrm{pr}_2 : M_r(\mathbf{R}) \times M_r(\mathbf{R}) \to M_r(\mathbf{R})$ by $\mathrm{pr}_1((a_{ij}),(b_{ij})) := (a_{ij})$, $\mathrm{pr}_2((a_{ij}),(b_{ij})) := (b_{ij})$. Then for all $A \in \mathscr{B}(M_r(\mathbf{R}))$,

$$\begin{aligned}(\mathrm{pr}_1)_* \pi_r(A) &= (\mathrm{pr}_1 \circ \xi_r)_* \mathscr{L}^{\otimes r}(A) \\ &= \mathscr{L}^{\otimes r}(\xi_r^{-1}(\mathrm{pr}_1^{-1}(A))) \\ &= \mathscr{L}^{\otimes r}(\xi_r^{-1}(A \times M_r(\mathbf{R}))) \\ &= \mathscr{L}^{\otimes r}(\{(s_1, \dots, s_r) \in (T_\varepsilon)^r \mid K_r^X(p_X(s_1), \dots, p_X(s_r)) \in A\}) \\ &= \mu_X^{\otimes r}(\{(x_1, \dots, x_r) \in X^r \mid K_r^X(x_1, \dots, x_r) \in A\}) \\ &= \underline{\mu}_r^X(A).\end{aligned}$$

This leads to $(\mathrm{pr}_1)_* \pi_r = \underline{\mu}_r^X$. In the same way, we get $(\mathrm{pr}_2)_* \pi_r = \underline{\mu}_r^Y$. Then π_r is an ε-transportation from $\underline{\mu}_r^X$ to $\underline{\mu}_r^Y$ and $\operatorname{def} \pi_r = 0 \le \varepsilon$. This means that $\mathrm{Tra}^{(M_r(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^Y) \le \varepsilon$. We get $d_{l_\infty\text{-Pr}}(X, Y) \le \underline{\square}_0(X, Y)$ by Theorem 2.9, the arbitrariness of ε, and Lemma 3.3.

Combining these two inequalities, we obtain the theorem. $\square$

Proof of Proposition 1.4. First, we prove $\underline{\square}_1(X_n, X) = n^{-1}$. For any parameter p_{X_n} of X_n, $\mathscr{L}(p_{X_n}^{-1}(\{p_1\})) = 1 - n^{-1}$. Then we have $d_{\{p_1,p_2\}}(p_{X_n}(s), p_{X_n}(s')) = 0$ for any $s, s' \in p_{X_n}^{-1}(\{p_1\})$. This means that $\underline{\square}_1(X_n, X) = n^{-1}$.

It is obvious that $\underline{\square}_0(X_n, X) = 1$ for any $n \geq 2$.

Next, we prove $d_{l_\infty\text{-Pr}}(X_n, X) = 1$ for any $n \geq 2$. We set the Borel set A_r on $M_r(\mathbf{R})$ by

$$A_r := \{(a_{ij}) \in M_r(\mathbf{R}) \mid \textit{there exist } k, l \in \{1, \dots, r\} \textit{ such that } a_{kl} = 1\}.$$

Then we have

$$\begin{aligned}
\underline{\mu}_r^{X_n}(A_r) &= \mu_{X_n}^{\otimes r}(\{(x_1, \dots, x_r) \in X_n^r \mid (x_1, \dots, x_r) \neq (p_i, \dots, p_i),\ i = 1, 2\} \\
&= \sum_{k=1}^{r-1} \binom{r}{k} n^{-k}(1 - n^{-1})^{r-k} \\
&= 1 - n^{-r} - (1 - n^{-1})^r.
\end{aligned}$$

It is obvious that $\underline{\mu}_r^X(A_r) = 0$ and $\underline{\mu}_r^X(B_{1-n^{-r}-(1-n^{-1})^r}(A_r)) = 0$ for any $n \in \mathbf{N}$. Let $\varepsilon > 0$. We have $\underline{\mu}_r^{X_n}(A_r) \leq \underline{\mu}_r^X(B_\varepsilon(A_r)) + \varepsilon$ if and only if $\varepsilon \geq 1 - n^{-r} - (1 - n^{-1})^r$. This means that $d_{\mathrm{Pr}}^{(M_r(\mathbf{R}), \|\cdot\|_\infty)}(\underline{\mu}_r^X, \underline{\mu}_r^{X_n}) \geq 1 - n^{-r} - (1 - n^{-1})^r$. For any $n \geq 2$, we get $d_{l_\infty\text{-Pr}}(X_n, X) = 1$ by Lemma 3.3. □

Acknowledgements

The author would like to thank Professor Takashi Shioya, Dr. Kazuki Ito, and Dr. Yu Kitabeppu for valuable discussions and many fruitful suggestions. He also thanks Professors Kei Funano and Takumi Yokota for their encouragement.

References

[1] V. I. Bogachev, Measure theory. Vol. I, II, Springer-Verlag, Berlin, 2007. MR2267655 (2008g:28002)

[2] Stewart N. Ethier and Thomas G. Kurtz, Markov processes, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. Characterization and convergence. MR838085 (88a:60130)

[3] K. Funano, Asymptotic behavior of mm-spaces, Doctoral Thesis, Tohoku University (2009).

[4] Andreas Greven, Pete Pfaffelhuber, and Anita Winter, Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees), Probab. Theory Related Fields **145** (2009), no. 1–2, 285–322, DOI 10.1007/s00440-008-0169-3. MR2520129 (2011c:60008)

[5] Misha Gromov, Metric structures for Riemannian and non-Riemannian spaces, Reprint of the 2001 English edition, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2007. Based on the 1981 French original; With appendices by M. Katz, P. Pansu and S. Semmes; Translated from the French by Sean Michael Bates. MR2307192 (2007k:53049)

[6] Takefumi Kondo, Probability distribution of metric measure spaces, Differential Geom. Appl. **22** (2005), no. 2, 121–130, DOI 10.1016/j.difgeo.2004.10.001. MR2122737 (2006a:53039)
[7] Wolfgang Löhr, Equivalence of Gromov-Prohorov- and Gromov's $\underline{\square}_\lambda$-metric on the space of metric measure spaces, Electron. Commun. Probab. **18** (2013), no. 17, 10. MR3037215
[8] A. M. Vershik, The universal Uryson space, Gromov's metric triples, and random metrics on the series of natural numbers, Uspekhi Mat. Nauk **53** (1998), no. 5(323), 57–64, DOI 10.1070/rm1998v053n05ABEH000069 (Russian); English transl., Russian Math. Surveys **53** (1998), no. 5, 921–928. MR1691182 (2000b:53055)
[9] Cédric Villani, Topics in optimal transportation, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR1964483 (2004e:90003)

Mathematical Institute
Tohoku University
Sendai 980-8578, Japan
E-mail: sb0m08@math.tohoku.ac.jp